DOCUMENT RESUME

ED 111 005 95 CS 202 221

AUTHOR Berdan, Robert
TITLE The Use of Linguistically Determined Groups in Sociolinguistic Research. Professional Paper 26.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO SWRL-PP-26
PUB DATE 73
NOTE 21p.

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Behavior Patterns; Caucasian Students; *Child Language; Comparative Analysis; Elementary Education; *Group Behavior; Interpersonal Relationship; *Language Research; *Language Usage; Negro Students; *Sociolinguistics

ABSTRACT

The established unit for reporting sociolinguistic data has been the sociologically determined group. Characteristically, only the mean rate of nonstandard usage for the group is reported for any linguistic feature. Such reporting obscures the possibility of linguistic heterogeneity within the group. Data from Los Angeles school children show that school classrooms are heterogeneous with respect to different forms of agreement usage. An alternative analysis is proposed in which children are grouped on the basis of linguistic criteria. Members of these groups share a common grammar. There is an implicational relationship among the grammars employed by such linguistically determined groups. The linguistically determined groups do not map exactly onto traditional sociologically determined groups. This relation may be expressed as the probability that a member of a sociologically determined group is a member of some linguistically determined group. This interpersonal variation may be distinguished from intrapersonal variation, or the probability that a rule of a particular grammar will apply. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 111605

SWRL

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Professional Paper 26

July 1973

THE USE OF LINGUISTICALLY DETERMINED GROUPS IN SOCIOLINGUISTIC RESEARCH

Robert Berdan

ABSTRACT

The established unit for reporting sociolinguistic data has been the sociologically determined group. Characteristically, only the mean rate of nonstandard usage for the group is reported for any linguistic feature. Such reporting obscures the possibility of linguistic heterogeneity within the group. Data from some published studies suggest that, for both Anglo and Black adults, sociologically determined groups are in fact linguistically heterogeneous. Data from Los Angeles school children show that school classrooms are heterogeneous with respect to different forms of agreement usage. An alternative analysis is proposed in which children are grouped on the basis of linguistic criteria. Members of these groups share a common grammar. There is an implicational relationship among the grammars employed by such linguistically determined groups. The linguistically determined groups do not map exactly onto traditional sociologically determined groups. This relation may be expressed as the probability that a member of a sociologically determined group is a member of some linguistically determined group. This interpersonal variation may be distinguished from intrapersonal variation, or the probability that a rule of a particular grammar will apply.

CS 202 281

THE USE OF LINGUISTICALLY DETERMINED GROUPS IN SOCIOLINGUISTIC RESEARCH*

Robert Berdan

The standard research design in sociol[illegible]tics for the past decade typically included (1) determination [illegible]et of sociological criteria which could reasonably be expected to differentiate linguistic behavior; (2) sampling the population according to these criteria; and (3) reporting differences in the way the resulting groups use each linguistic feature of interest.

This was the design employed in much of Labov's classic study of English in New York City (Labov, 1966). A previous sociological survey provided information for grouping informants according to age, sex, ethnicity, and a composite scale of socioeconomic class. Differences in linguistic behavior were found to correlate with groups defined by the intersections of these parameters.

This same paradigm of observing the linguistic behavior of sociolocically determined groups has been utilized in most recent studies of Black English, e.g., the Detroit study under Shuy (Shuy, Wolfram, & Riley, 1967) and the resulting reporting by Wolfram (1969); the 1968 work by Labov and his associates (Labov et al., 1968); and, the recent study by Fasold (1972) in Washington, D. C.

Membership in the sociologically determined groups used by Wolfram was established by assigning weighted ratings to a person's educational level, his occupation, and the percentage of houses in his block that had all their plumbing, with adjustments for the median income of the census tract. Table 1 shows the differences among these groups with respect to one linguistic feature: multiple negation.

This design is admirably suited to Labov's original purpose: to demonstrate that linguistic behavior is differentially conditioned by social stratification. He made his case well. Today, no one disagrees that linguistic behavior is associated with other markers of social stratification. Current work seems to be dedicated to filling in the gaps, showing which linguistic features correlate with which sociological criteria.

In the past decade another purpose has evolved in sociolinguistic research: the writing of grammars. The term "grammar" is used here in the restricted sense of the transformational grammarian. An increasing number of sociolinguists are attempting to relate that formal model of language to empirical observations of the linguistic community.

---

*Paper presented at SECOL IX, April 21, 1973, Charlottesville, Virginia.

Table 1

Percentage of Realized Multiple Negation
by Sociologically Determined Group

Social Group	Mean	Median
Upper Middle White	1.2	0.0
Upper Middle Negro	8.2	0.0
Lower Middle Negro	12.3	0.0
Upper Working Negro	54.7	53.5
Lower Working Negro	77.8	81.3

Source: Wolfram, 1969, p. 156.

The purpose of sociolinguistic research has in this way changed, but the research design has remained constant. This paper argues that the research design which served Labov's original purpose so well needs serious modification for the current sociolinguist attempting to construct grammars to account for linguistic behavior.

There are several problems which result from attempting to construct grammars with data derived from sociologically determined groups. In large part these problems result from the nature of sociolinguistic variation. Data on language variation are explicitly ignored by many theoretical linguists, but are the primary focus of the sociolinguist. The papers already mentioned document convincingly two kinds of variation: variation within the speaker and variation between groups of speakers. Any given individual may use different surface realizations of some particular linguistic item at different points in time. Labov's (1969) article on copula deletion is the classic argument. The data in Table 1 are a typical example of the variation found between groups.

There is another kind of variation that is mentioned chiefly only in footnotes. That is the linguistic variation among the individuals within a single sociologically determined group. The standard unit of reporting in studies of Black English is the mean percentage of non-standard usage of the sociologically determined group. It is tacitly assumed that this mean, which is descriptive of the group, may be considered an attribute of any member of the group. Groups are different, but members of a single group are all assumed to be the same. If pressed on this point no sociolinguist who has actually confronted data will assert that sociologically determined groups are necessarily linguistically homogeneous, but he conducts his research as though the myth were fact.

Consider again the data in Table 1. Wolfram is somewhat atypical among sociolinguists in that he not only reported the mean, or average, but also reported the median, or middle ranked score. The means show an ordered progression of increasing percentages of nonstandard negation correlating with decreasing social status. The medians show a somewhat different picture. In each of the three middle class groups at least half of the informants categorically did not use nonstandard negation.

Given that at least half the informants in each of these groups used no nonstandard negation it is obviously not the case that all members of UMW, UMN, and LMN groups used 1.2%, 8.2% and 12.3% nonstandard, respectively. And, given that there are 12 members in each group, it cannot possibly be the case, for the first two groups at least, that some one person categorically did use nonstandard negation. The only conclusion must be that not all individuals sometimes used nonstandard negation.

The use of sociologically determined groups results in the confounding of interpersonal variation with intrapersonal variation. From the medians in Table 1 it is possible to determine that the groups are not linguistically homogeneous, but from those figures alone it is not possible to determine how many speakers used what proportion of nonstandard responses.

Wolfram does, however, provide more information about the nonstandard negation used by these groups. In UMW there was in fact only one individual who used nonstandard negation. Simple arithmetic shows that he used it 14.4% of the time. In UMN and LMN there were four and five individuals, respectively, who used nonstandard negation. Only two informants in LWN categorically used nonstandard negation in all syntactic environments.

There are several proposals in the literature for incorporating figures such as the means in Table 1 into grammars (e.g., Labov et al. 1968; Fasold, 1972; Houston, 1972). Proposals differ in detail, but in general they would associate probability functions with the rules that derive, in this instance, multiple negation. For speakers classed as UMW that probability function would have a value of .012, etc. Of the five grammars that result from these data there would be only grammars with variable rules of nonstandard negation; no grammar with a categorical rule of nonstandard negation, and no grammar with a categorical rule of standard negation. This despite the fact that 26 out of the 48 informants--over half--demonstrated categorical standard negation. Clearly, some generalizations are being missed in a system which writes grammars that no one speaks and assigns persons with identical language behavior to three different grammars.

If this were a unique instance of a non-homogeneous sociologically determined group one might safely ignore it and continue to assume that such groups are, in general, linguistically homogeneous. It is difficult to determine within-group variance from many published studies. The

standard statistic employed by most sociolinguists is the mean. There is no report of the range of scores, the standard deviation, nor any other statistic which might show that a group could be heterogeneous. However, there are some instances in which data are published for individuals as well as for groups, and from these it appears that within-group variation may be the rule, not the exception.

The figures in Table 2 are from Labov's (1966) famous "fourth floor" experiment. They represent the percentage of employees using constricted /r/ in each of three New York department stores. These figures are for what Labov described as "The most homogeneous sub-group in the three stores: native New York, white, saleswomen."

Table 2

Percentage of Native New York White Saleswomen Using /r/

Realization	Store		
	Saks	Macy's	Klein's
Categorical Non-Constricted /r/	33	41	70
Variable Constricted Non-Constricted /r/	34	31	26
Categorical Constricted /r/	33	28	4

Source: Labov, 1966, p. 79.

These individuals are grouped into three sociologically determined groups. Sex and ethnicity are constants, while place of employment is allowed to vary. Following the research paradigm which uses sociologically determined groups, one might posit three dialects: a Saks grammar, a Macy's grammar, and a Klein's grammar, with respectively greater probabilities of /r/ deletion. Notice again that there will be no grammar with categorical constricted /r/ nor any grammar with categorical non-constricted /r/, despite the fact that over two-thirds of the individuals evidenced categorical grammars.[1]

[1]The proportion of categorical grammars would doubtless decrease if more instances were elicited from each informant. The point remains that no store grammar would accurately predict the linguistic behavior of a large number of the individuals.

Alternatively, one could construct three grammars:

Grammar I: no rule of /r/-vocalization

Grammar II: variable rule of /r/-vocalization

Grammar III: obligatory rule of /r/-vocalization

Now the individuals in Table 2 can be grouped linguistically, according to the grammar they employ. This allows another interpretation of the figures in Table 2. These figures now may be understood as probabilities that, given a person's place of employment he belongs to a certain linguistically determined group. That is, if a woman works at Saks, the probability that she uses Grammar I is p = .33. For an individual who uses Grammar I, there is no question of /r/-vocalization.

The probability that a saleswoman working at Klein's uses Grammar II is p = .26. Once it is determined that an individual uses Grammar II the relative frequency with which the /r/-vocalization rule applies remains to be determined. For that, data different from these summarized in Table 2 are needed.

If grammars are to achieve even observational adequacy it seems that the unit of investigation and reporting must be the linguistically determined group, rather than the sociologically determined group. The sociologically determined group is constructed of individuals judged similar on a set of sociological parameters. The linguistically determined group is constructed of individuals who evidence similar linguistic behavior. Only the latter allows separation of linguistic variation within the individual from variation between individuals. This distinction may be stated alternatively as the distinction between the probability that an individual employs a given grammar, and the probability that a particular rule within the grammar will apply.

DeCamp (1971) argued for a slightly different reason against using the sociologically determined group as the basic unit of linguistic reporting. He observed that many sociological criteria are continuous rather than discrete. When a sociolinguistic study reports the median income for each group of informants it does not follow that each informant in the group had that exact income; it may frequently be the case that the lowest individual in one group and the highest individual in the next group down the ladder are more similar than are two arbitrarily chosen individuals from one group. Demonstrating that two groups are different does not imply that either group is internally homogeneous.

It is also not the case that sociological parameters inter-correlate perfectly. If they did, any one parameter would provide as much information as any other. If such parameters as housing, occupation and education do not correlate perfectly, it is logically impossible for

linguistic behavior to correlate perfectly with each of them. The extent to which linguistic behavior correlates with the composite scales used by sociolinguists is a hypothesis which can be tested. But it cannot be tested by reporting the rate of mean nonstandard response for sociologically determined groups. Such observations test a weaker hypothesis: does linguistic behavior show any correlation with sociologically determined groups. Neither of these hypotheses makes any claims about causal relationships between sociological status and linguistic behavior.

Another argument against the use of sociologically determined groups is the existence of what Carden (1972) has called "randomly distributed dialects." Carden has examined a number of linguistic features that vary from individual to individual but do not appear to correlate in any way with sociological parameters. He attributes this to the possibility these aspects of the grammar are under-determined by the data available to the language learner. Given the rarity in natural language of the esoteric sentences by which some transformational rules rise and fall, it is not unlikely that the child acquiring his language never heard the crucial data. Existing theories of language acquisition implicitly predict that there will be at least some randomly distributed dialects. If these features are in fact randomly distributed across social groups, then the effect of analysis by sociologically determined group will be to submerge the nature of this variation.

The only way to determine whether some new feature to be studied is randomly distributed or is correlated with sociological parameters is to group individuals into linguistically determined groups on the basis of use of the feature. These groups may then be compared on sociological parameters.

One study which separates interpersonal variation from intrapersonal is Bickerton's (1971) treatment of infinitive markers in Guyanan Creole. Bickerton observed that some of his informants, like Labov's New Yorkers, categorically used one marker, some categorically used the other marker, and some demonstrated a limited set of patterns of variation between the two markers. However, the small number of informants, distributed across six grammars, precludes any statements relating grammars to sociological parameters.

Most of the current work on variation comes from the study of Black English, and the most intensively studied aspect is undoubtedly the -Z, -D inflectional morphology. The following data are from children in three Los Angeles neighborhoods: a low income Black neighborhood (LB), a middle income Black neighborhood (MB), and a low income Anglo neighborhood (LA).

The figures in Table 3 from the 1970 census give some sociological information about the three neighborhoods. The low middle income Black neighborhoods were clearly different, particularly with respect to income distributions.

Table 3

Sociological Characteristics of the School Neighborhoods

Sociological Characteristics	Low Income Black	Middle Income Black	Low Income Anglo
% Negro	77.4	62.8	.01
% High school grads (over 25 years old)	63.4	80.1	42.1
% Unemployed (male, over 16)	9.2	3.7	11.1
% Families below poverty level	14.8	3.4	10.4
Median family income	$ 8,804	$15,077	$ 9,517
Median house value	$20,800	$38,500	$21,800

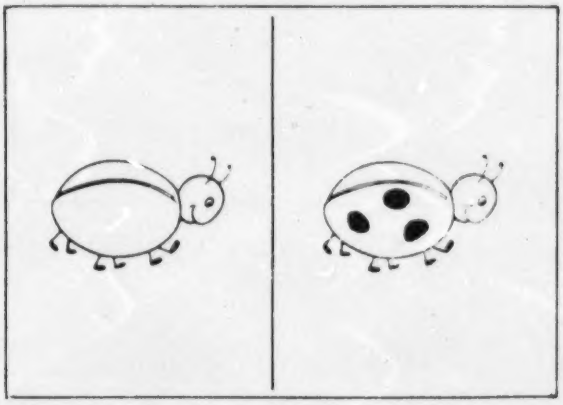
Source: 1970 Los Angeles-Long Beach SMSA Census Tract Information.

All the children in one first grade and one sixth grade in each neighborhood were interviewed individually three times. The schools were considerably less integrated than the census figures would indicate; there was one Anglo child in each of the four classrooms in the Black neighborhoods.

The setting for the interviews was not unlike a classroom. A set of structured tasks was presented to the child by an adult interviewer of the same ethnicity. The tasks usually required the child to respond to some question about a picture.[2] It was possible to constrain the question and the picture in such a way that there was high probability the child's response would contain the desired linguistic construction, without presenting any dialect-sensitive model for him to imitate. This increased the comparability of responses across informants and allowed the elicitation of multiple instances of the constructions in controlled environments. An item from one task designed to elicit have, do, and negation, along with sample responses, is shown in Figure 1.

[2]The tasks used in the first set of interviews were comparable to those reported in Berdan and Pfaff (1972). Later interviews used similar tasks designed to elicit other linguistic features.

Fig. 1. Sample elicitation item and responses.



Interviewer: What's the difference between this bug and that bug?

Child: This bug {doesn't have / don't have / hasn't got / don't got} {any / no} spots and this one {does / has / do}

Responses to the first round of interviews were tabulated in the linguistic categories listed in Table 4. The data represent a relatively careful style. This is not the vernacular of peer group casual conversation. However, it would be a mistake to equate these data with Houston's (1969) characterization of "School Register." The children obviously enjoy performing the tasks; they receive repeated affirmative feedback for successfully performing what they view as the intent of the task. There is no wordlist intonation; utterances are longer than the minimum required to perform the task. In short, there is no reason to believe that the responses do not represent natural, though careful, language.

Table 5 gives the mean proportion of nonstandard response for each feature by classroom.[3] These classrooms may be viewed as sociologically determined groups. Each group contains only children of the same ethnicity, of the same age, and from the same neighborhood. Unlike the groups in studies such as Wolfram (1969) and Fasold (1972), occupation and education of parents are not considered. However, unlike the groups

[3]Excluding Anglo children in the classrooms in Black neighborhoods, and Mexican-American children who could not perform the tasks in English in the Anglo neighborhood.

Table 4

Standard and Nonstandard Responses to Linguistic Features

Features	Standard Responses	Nonstandard Responses
GOT (main verb)	any use with auxiliary have	used without auxiliary have
/ð/ (initial)	fricative /ð/	affricate /dð/ or stop /d/
DO NEG	doesn't[3rd sing]	don't[3rd sing]
DO	does[3rd sing]	do[3rd sing]
HAVE	has[3rd sing]	have[3rd sing]
M.V. (regular main verb agreement)	cook [s] ride [z] catch[ɨz]	cook [Ø] ride [Ø] catch[Ø]
Negation (verb & indefinite object)	doesn't have any	don't have no
IS (copula)	This one is red	This one Ø red

Table 5

Rate of Nonstandard Usage for Each Feature by Classroom

Feature	Low Income Black				Middle Income Black				Low Income Anglo				Overall Mean
	1st Grade		6th Grade		1st Grade		6th Grade		1st Grade		6th Grade		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
GOT	.94	.24	.97	.05	1.00	.00	.00	.00	.94	.19	.35	.43	.70
/d/	.65	.30	.56	.31	.60	.30	.40	.32	.34	.37	.33	.32	.46
DONT	.94	.17	.35	.45	.28	.43	.02	.06	.57	.47	.42	.45	.43
HAVE	.83	.29	.37	.43	.21	.33	.05	.13	.15	.32	.07	.25	.28
M.V.	.68	.29	.26	.28	.26	.28	.12	.28	.12	.23	.08	.23	.25
DO	.96	.16	.30	.46	.12	.31	.00	.00	.10	.28	.00	.00	.25
NEG	.66	.42	.11	.28	.28	.36	.02	.09	.31	.35	.02	.07	.23
IS	.21	.31	.10	.14	.04	.10	.01	.0?	.03	.10	.02	.10	.07
Mean	.73		.38		.35		.08		.32		.15		.33

in those studies, these groups are actually linguistic communities in the sense that there is primary interaction among the members. Members of these groups form, in large part, the peer groups which are universally acknowledged to influence children's speech. School classrooms are in fact the sociologically determined groups used in a number of sociolinguistic studies, e.g., Garvey & Dickstein, (1970); Baratz, (1969); DeStefano, (1970).

Several generalizations may be observed from the means in Table 5. Without exception, for every feature, in each school, the mean rate of nonstandard response was lower for the sixth grade than for the first grade. Also, for every feature, for both first and sixth grade, the rate of nonstandard response is lower in the middle income Black school than in the lower income Black school. With one exception, the rate of nonstandard response for the Anglo first grade is lower on every feature than the rate for either Black first grade. In general, the middle income Black sixth grade appears less nonstandard than the Anglo sixth grade.

These suggestions that age, income level, and ethnicity are relevant factors in describing dialect are scarcely novel. The only findings that may be at all surprising are the number of instances in which middle income Black children appear less nonstandard than low income Anglo children.

Such observations are quite appropriate to the linguistic characterization of social groups, but are of limited utility for the construction of grammars. The reason is apparent on examination of the standard deviations in Table 5. For most features the groups are linguistically quite heterogeneous. Scores of individuals rarely show normal distributions. A few features have distributions like that shown for /d/ in Figure 2. Most, however, have bimodal distributions like that shown for DO NEG in Figure 3. Differences in means represent chiefly differences in the proportions of children who categorically did and categorically did not use a particular feature.

The means in Table 5 could lead to predictions such as, the probability of don't rather than doesn't for children in the middle income Black first grade is .30. However, Figure 3 shows that p = .30 predicts the behavior of exactly one child in the classroom. A grammar which incorporates such a probability function is a statement that the use agreement with negative do is variable. However, the majority of individuals for whom this grammar is posited categorically used doesn't or categorically used don't. Not only is the value of the probability function meaningless for most individuals, the use of probable rather than categorical agreement is observationally inadequate.

The alternative which has already been suggested is to group individuals on the basis of linguistic behavior. Several methods of grouping could be employed; the traditional three-way distinction be-

Fig. 2. Distribution of /ɛ/ scores for middle income Black 1st grade.

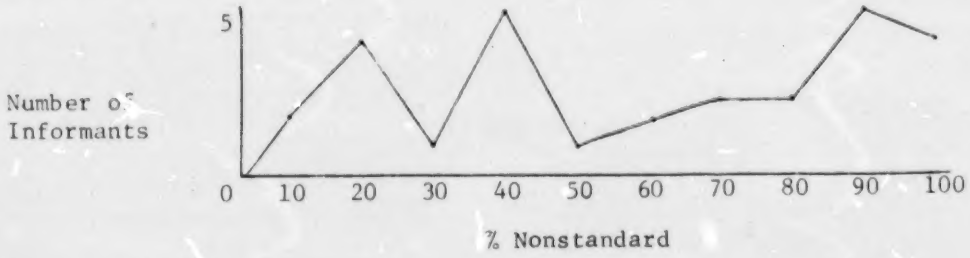
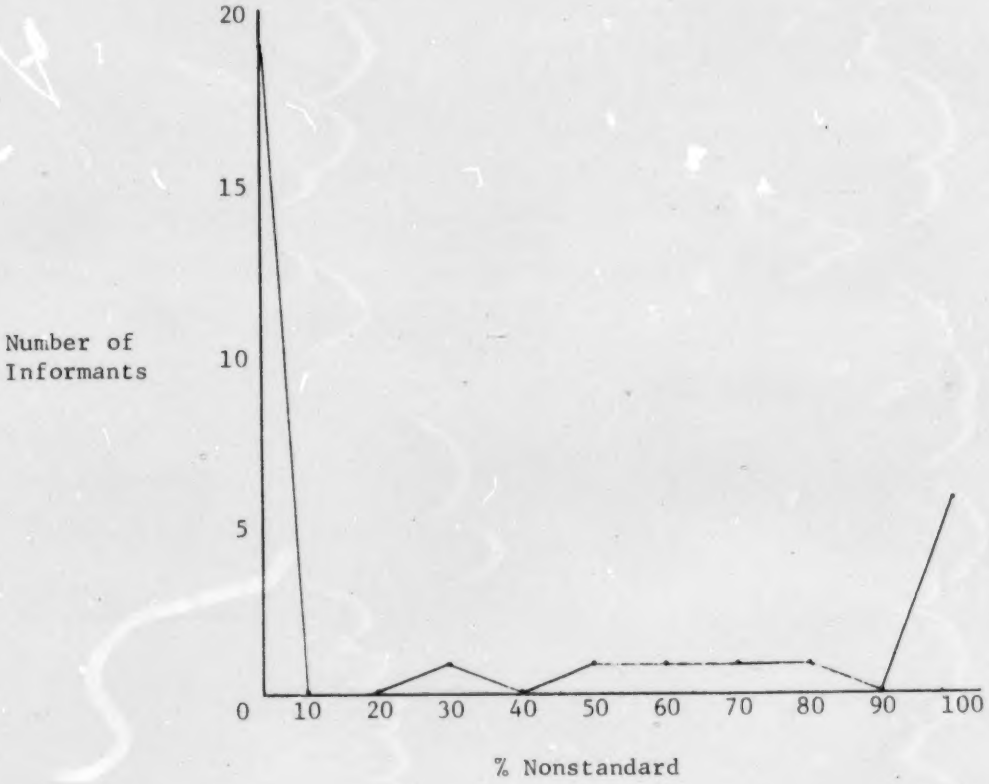


Fig. 3. Distribution of DO NEG scores for middle income Black 1st grade.



tween categorical use (1), categorical non-use, (0) and variation (X) is used here. Considering only the four types of agreement included in Table 5 (affirmative do, negative do, have, and regular main verbs) there are 81 logically possible combinations of 1, 0, and X. If these four uses of agreement were unrelated and randomly distributed through the populations, each of the 81 possible grammars should have equal probability of occurence. In fact, ten of the possible combinations, or grammars, account for 81 percent of the individuals who used all four of the constructions.[4] The rest of the individuals are scattered, one per grammar.[5] Table 6 show the distribution by classroom for the ten non-idiosyncratic grammars.

Table 6

Distribution of Informants by Classroom and Grammar

Grammars		G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9	G_{10}
Features	HAVE	0	0	1	X	0	0	0	X	0	X
	DO	0	0	1	0	0	0	0	1	0	0
	DONT	0	0	1	0	X	1	1	1	X	1
	M.V.	0	X	X	X	0	0	X	X	X	X
Classrooms	LB-1	0	0	8	0	0	0	0	3	0	0
	LB-6	5	3	2	0	0	0	0	0	0	1
	MB-1	4	8	0	5	0	0	0	0	1	2
	MB-6	11	2	0	1	0	0	0	0	1	0
	LA-1	4	3	1	2	2	2	3	1	0	0
	LA-6	9	1	0	0	4	4	2	0	1	0
Total		33	17	11	10	6	6	5	4	3	3

[4]Not all of the children used all three of the lexical items HAVE, DO and DONT. In particular, some children consistently used got rather than has or have. There were also a few children who did not use affirmative do or does. These children have been excluded from the calculation on non-idiosyncratic grammars. However, 88% of these individuals are consistent with the ten grammars in Table 6.

[5]It is impossible to evaluate the significance of idiosyncratic performances, particularly given the possibilities of measurement error and the exigencies of performance. It seems best to set aside these idiosyncratic performances until there are more data to evaluate them.

From the figures in Table 6 it is possible to make certain sociolinguistic and linguistic generalizations not apparent from Table 5. Table 5 shows no classroom with means of 0.00 or 1.00, i.e., categorical usage, for all four features. Table 6 shows that the categorically standard grammar was the most highly represented grammar. On the other hand, only one child in the entire study had variable responses on all four features.

Not every grammar was used by children in every classroom. No Black children used grammars G_5, G_6, or G_7. A maximum of one Anglo child used each of the grammars G_3, G_8, G_9, or G_{10}. The other grammars are used by both Anglo and Black children.

These facts are reflected in the reorderings of grammars in Tables 7 and 8. In Table 7 it appears that the use of agreement with DO is categorical for all Black informants with non-idiosyncratic grammars. This is also true of DONT for all but two informants. Agreement with regular main verbs, on the other hand, is used variably in all except the categorically standard grammar.

Table 7

Distribution of Black Informants by Grammar

Grammars		G_1	G_2	G_4	G_9	G_{10}	G_8	G_3
Features	DO	0	0	0	0	0	1	1
	DONT	0	0	0	X	1	1	1
	HAVE	0	0	X	0	X	X	1
	M.V.	0	X	X	X	X	X	X
No. of Informants		20	13	9	2	3	3	10

Table 8

Distribution of Anglo Informants by Grammar

Grammars		G_1	G_2	G_5	G_6	G_7	G_4
Features	DO	0	0	0	0	0	0
	HAVE	0	0	0	0	0	X
	M.V.	0	X	0	0	X	X
	DONT	0	0	X	1	1	0
No. of Informants		13	4	6	6	3	2

For Anglo children, only DONT categorically lacks agreement, as in G_6 and G_7. Affirmative DO, however, is categorically standard for all but two Anglo children with non-idiosyncratic grammars.

Using the observed frequencies of grammars reported in Table 6 it is possible to hypothesize probabilities that a child from a particular classroom, or sociologically determined group, will be a member of a particular linguistically determined group. Table 9 gives probabilities for each of the grammars in Table 7, broken down by classroom. It is apparent that the right-most, or most nonstandard grammars are most probable for children from the low income Black first grade, while the left-most or most standard grammars are most probable for children from the middle income Black sixth grade. With respect to these features, these two sociologically determined groups are linguistically distinct; they share no grammars. The probabilities for the middle income first grade and the lower income sixth grade are, not surprisingly, less distinctive. Low income and young age appear to increase the probability of nonstandard grammars; greater income or increased age increase the probability of more standard grammars.

Of the 81 possible grammars which the children could have used, those actually used by more than one Black child show a distinct pattern. DeCamp (1971, 1972) has hypothesized that the grammars employed in any speech community can be ordered linearly and displayed on a Guttman scalogram (Torgerson, 1957). With a dichotomy between categorically standard (0) and not categorically standard (X or 1) shown as a stepped line in Table 7, these data form an implicational array with a high coefficient of reproducability (R). The seven grammars represent 60 individuals, or a table with 240 cells. Only two cells in G_9 are anomolous, with a resulting $R = 0.99$.

Table 9

Probability of Grammars for Black Children, by Classroom

Classrooms	Grammars G_1	G_2	G_4	G_9	G_{10}	G_8	G_3
LB-1	.00	.00	.00	.00	.00	.27	.73
MB-1	.20	.40	.25	.05	.10	.00	.00
LB-6	.45	.27	.18	.00	.09	.00	.18
MB-6	.73	.13	.07	.00	.07	.00	.00

The separation of categorical usage from variable usage, by individual, results in groups which can be judged in well-defined linguistic terms to be homogeneous. Among the individuals who employ grammars G_1 or G_6 there remains no question of intrapersonal variation. All use each feature categorically. This lack of intrapersonal variation is not necessarily true of the other grammars. $\underline{X}$ is defined as a percentage ranging from less than one to greater than zero. This means that persons who use agreement with regular main verbs one time in ten are lumped together in G_2 with persons who do so nine times in ten. One might justifiably question whether such individuals are in fact linguistically similar. In this particular case they definitely are not.

There appear to be two kinds of variable usage of main verb agreement. There are a large number of children who use agreement with all verbs except those which end in strident consonants, where the agreement morpheme is syllabic [ɨz]. This is true of all but four of the 17 children in G_2. These four children show variation in use of the agreement morpheme without respect to phonological conditioning. The other children appear to have categorical agreement, but lack an epenthesis rule (Hoard & Sloat, 1971).

Data from the second and third round of interviews with these same children support this hypothesis. The same children who use agreement with all verbs except those ending in strident consonants appear to use the plural with all nouns except those which end in strident consonants. The same is true of possessive. Such observations cannot be made by using only sociologically determined groups. The mean proportion of epenthesis for the classrooms is always less than one but greater than zero. This precludes determining whether the same individuals who do not show epenthesis with verbs also do not use epenthesis with noun morphology. However, grouping individuals who are linguistically similar makes such generalizations possible.

The use of linguistically determined groups in sociolinguistic research does not automatically result in meaningful grammars. But it is the only basis from which meaningful sociolinguistic grammars may be constructed. This is not to say that one looks at linguistic behavior totally without regard to sociological information. To group all individuals in the world who use word-final devoicing into a single linguistically determined group for writing grammars is as absurd as grouping all males in the world into a single sociologically determined group.

Sociologically determined groups are only relatively sociologically homogeneous. The probability that they are linguistically homogeneous is too small to make them useful as the basis for construction grammars. Given the linguistic heterogeneity of such groups it is impossible to separate interpersonal variation from intrapersonal variation. They also preclude the observation of regular use of rules across features by individuals.

Patterns that are both sociologically and linguistically interesting emerge only as one focuses on groups, rather than on individuals. There are few perceptible patterns in the total array of idiosyncratic performances exhibited by the children in the study. However, for the Black children at least, there is an implicational pattern among non-idiosyncratic grammars. Children who used nonstandard do also used nonstandard don't; those who used nonstandard don't also used nonstandard have. Children who used any of these nonstandard lexical items, also showed nonstandard agreement with regular verbs.

The dichotomy drawn between categorical and variable use of linguistic features is not the only, nor necessarily even the best, determinant of homogeneity. It is one dichotomy which is linguistically well-defined. Much of what appeared in Table 5 to be inherent variation is in fact interpersonal variation. The literature of inherent variation is much too rich to suggest that it is all an artifact of the sociologically determined group. However, separation of categorical usage from variable usage through linguistically determined groups will provide a more reliable foundation for studies of inherent variation.

In his New York study, Labov observed that there are two approaches to the study of social variation in language. They are essentially the dichotomy drawn here between the sociologically determined group and the linguistically determined group. The approach selected for the New York study was to "consider various sections of the population, and determine the values of the linguistic variables for each [social] group." But Labov also states that, "When we have finished this type of analysis, we may turn to the second approach, and use the concept of linguistic class as a first step towards establishing the over-all structure of New York City English."

Sociolinguistic analysis by sociologically determined group is definitely not yet finished. However, if sociolinguists are to construct meaningful grammars the time has come also to use linguistically determined groups.

REFERENCES

Baratz, Joan C. A bi-dialectal task for determining language proficiency in economically disadvantaged Negro children. Child Development. 1969, 40, 889-901.

Berdan, Robert & Pfaff, Carol W. Sociolinguistic variation in the speech of young children: an experimental study. Professional Paper 21, 1972. SWRL Educational Research and Development, Los Alamitos, California.

Bickerton, Derek. Inherent variability and variable rules. Foundations of Language. 1971, 7, 457-492.

Carden, Guy. Dialect variation and abstract syntax. In The Georgetown Linguistics Forum, ed. Roger W. Shuy. Washington, D.C.: Georgetown University Press, 1972.

DeCamp, David. Implicational scales and sociolinguistic linearity. Linguistics. 1971, 73, 30-43.

DeCamp, David. What do implicational scales imply? Paper read at the SECOL VIII, Washington, D.C., 1972.

DeStefano, Johanna S. A sociolinguistic investigation of the productive acquisition of a school language instruction register by black children. Ph.D. dissertation, Stanford University, 1970.

Fasold, Ralph W. Tense marking in Black English. Washington, D.C.: Center for Applied Linguistics, 1972.

Garvey, Catherine J. & Dickstein, Ellen. Levels of analysis and social class differences in language. Report No. 83. Center for Social Organization of Schools, Johns Hopkins University, Baltimore, Maryland, 1970.

Hoard, James & Sloat, Clarence. The inflectional morphology of English. Glossa. 1971, 5, 47-56.

Houston, Susan H. A sociolinguistic consideration of the Black English of children in northern Florida. Language. 1969, 45, 599-607.

Houston, Susan H. Contingency grammar: introduction to a general theory of competence and performance. Papers in Linguistics. 1972, 5, 10-27.

Labov, William. Contraction, deletion, and inherent variability of the English copula. Language. 1969, 45, 715-762.

Labov, William, Paul Cohen, Clarence Robins, & John Lewis. A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Cooperative Research Project No. 3288, 2 vols, 1968. Washington, D.C.: Office of Education.

Shuy, Roger W., Walter A. Wolfram, & William K. Riley. Linguistic correlates of social stratification in Detroit speech. Final Report, Project 6-1347, 1968. Washington, D.C.: U.S. Office of Education.

Torgerson, Warren. Theory and method of scaling. New York: John Wiley & Sons, 1957.

Wolfram, Walter A. A sociolinguistic description of Detroit Negro speech. Washington, D.C.: Center for Applied Linguistics, 1969.